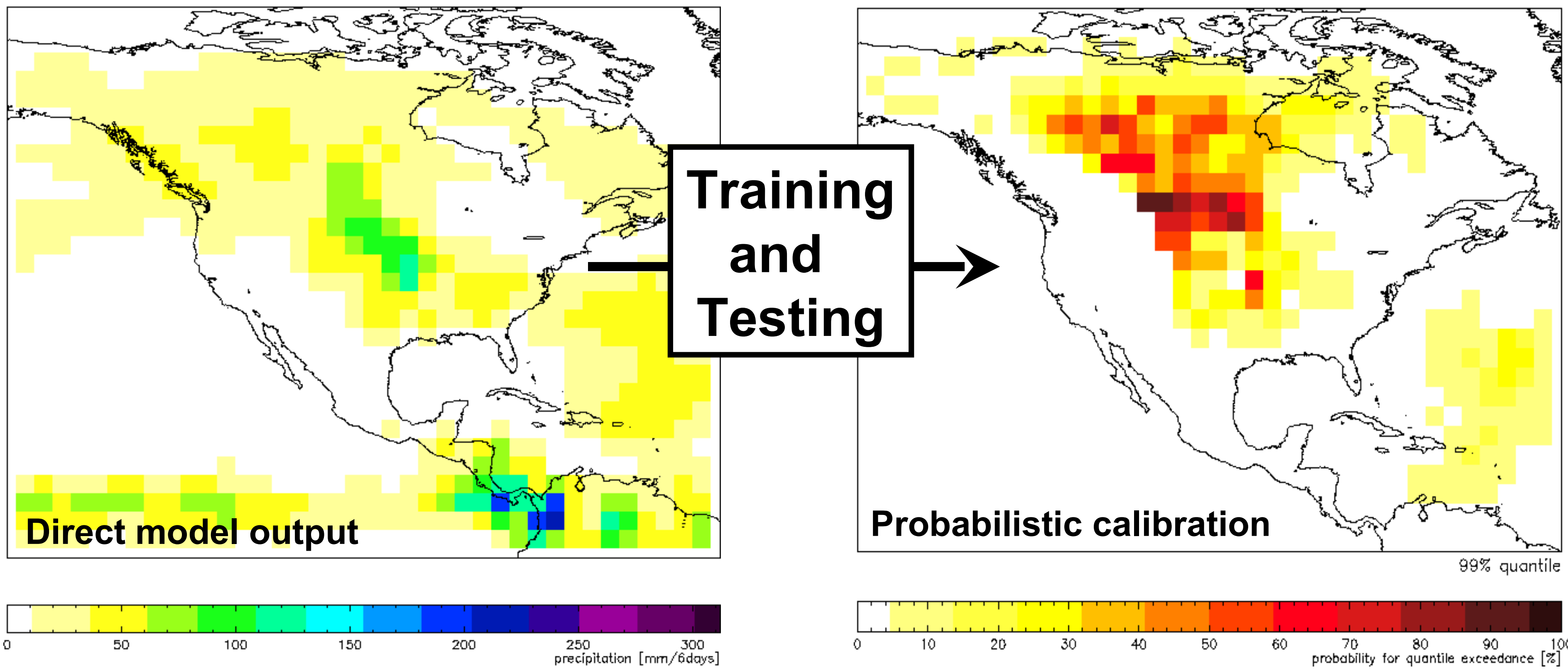
Verification of probabilistic calibrations for deterministic GFS precipitation forecasts



Johannes Jenkner and **Simon Mason**
International Research Institute for Climate and Society

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Introduction of datasets

NOAA Earth System Research Lab.
(Re-)forecasts

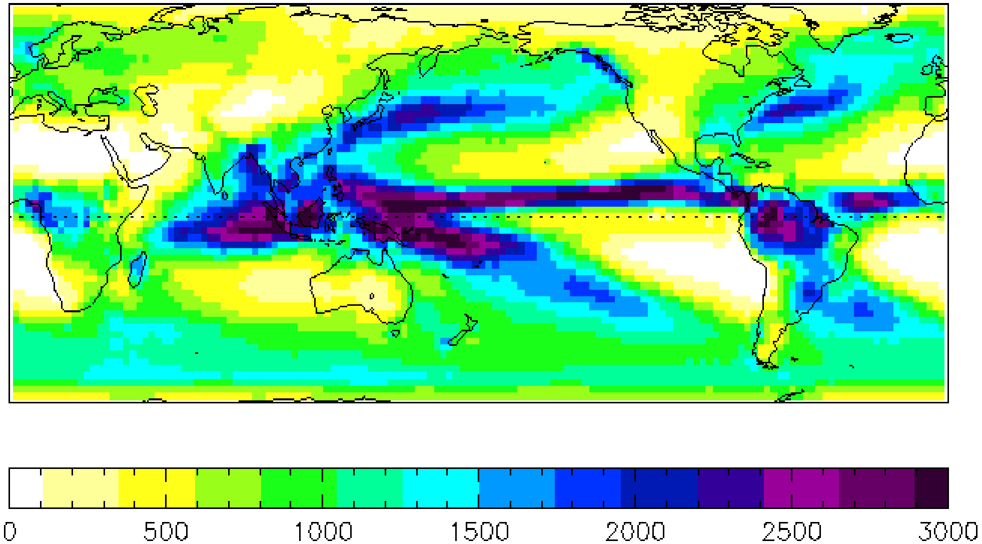
- GFS version based on 1998 physics
- T62 resolution with 28 vertical σ - levels
- Daily 15 member ensemble from 1979 to present
- 0000 UTC initial conditions from NCEP-NCAR reanalysis

Focus on forecast rainfall sum for a week (6 forecast days) and heavy events (99% quantile, exceeded once in 100 days)

GPCP One-Degree Daily
Multisatellite Rainfall Estimates

- Threshold-Matched Precipitation Index at low latitutes
- Rescaling of TOVS satellite measurements at high latitudes

Annual rainfall climatology [mm/year]



Data coverage Oct 1996 - Apr 2008: 10.6 years for training and 1 year for testing

Training: Oct 1996 - Apr 2007 Testing: May 2007 - Apr 2008

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Verification of direct model output

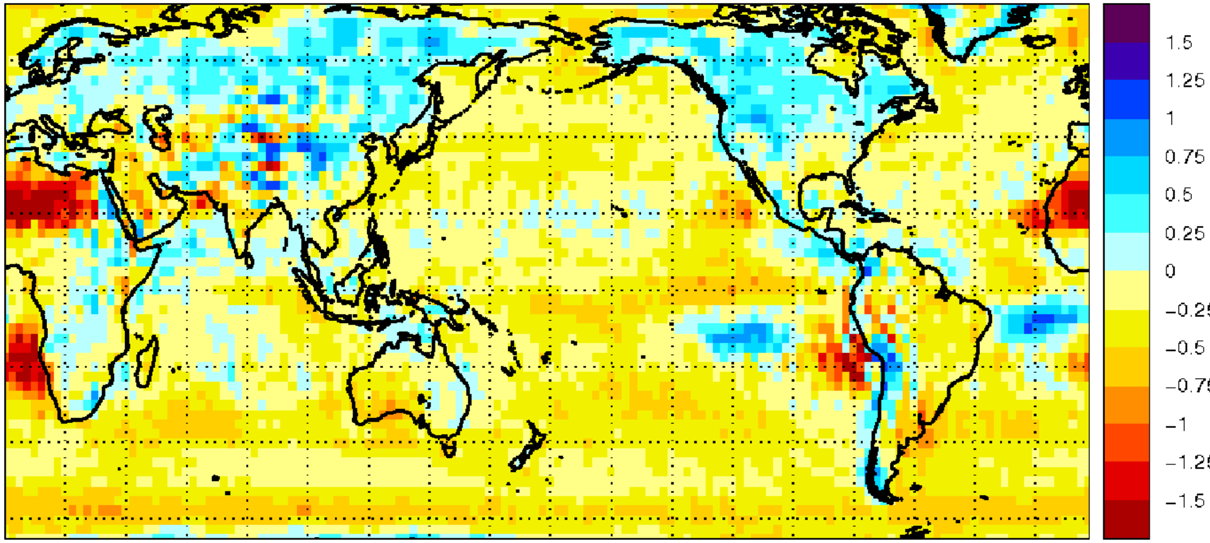
p=99%

Ensemble mean

Bias

$$\mathrm{QD}'(p) = 2\,\frac{q_{mod}(p) - q_{obs}(p)}{q_{mod}(p) + q_{obs}(p)}$$

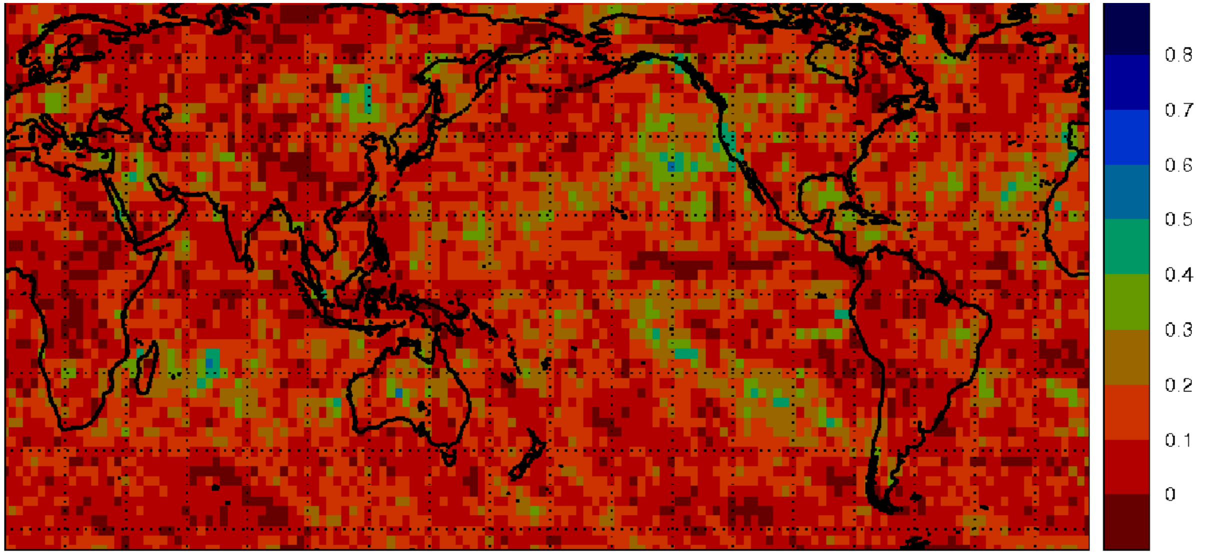
(spatial mean: -0.20)



Potential skill

$$\mathrm{PSS}(p) = 1 - \frac{Misses(p)}{Misses_{rand}(p)}$$

(spatial mean: 0.11)



→ **Jenkner et al. (2008)**

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Probabilistic calibration

- **Ensemble mean** of ESRL forecasts
- Principal com-ponents based on **sum of squares**
- Regression with a resolution of **2.5°x 2.5°**

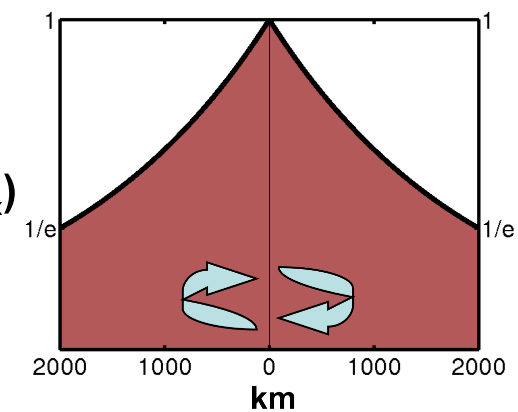
Input:
Principal component scores based on rainfall pixels within r_{max}

Logistic Regression

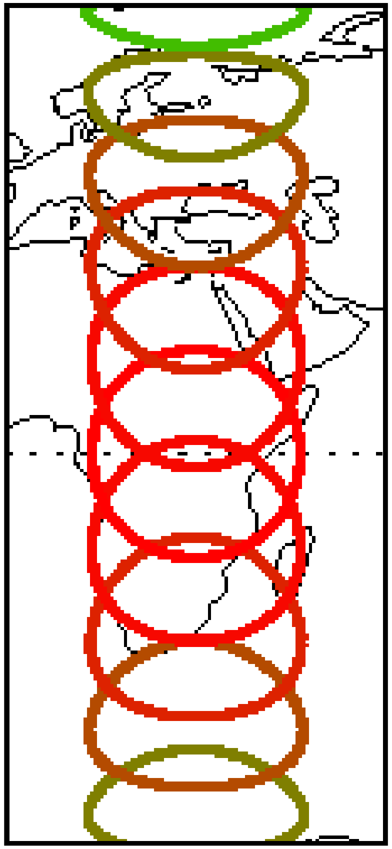
Output:
Probabilities for exceedance of rainfall frequency threshold

How many principal components are to be included into the regression?

weight = $\exp(-r/r_{max})$



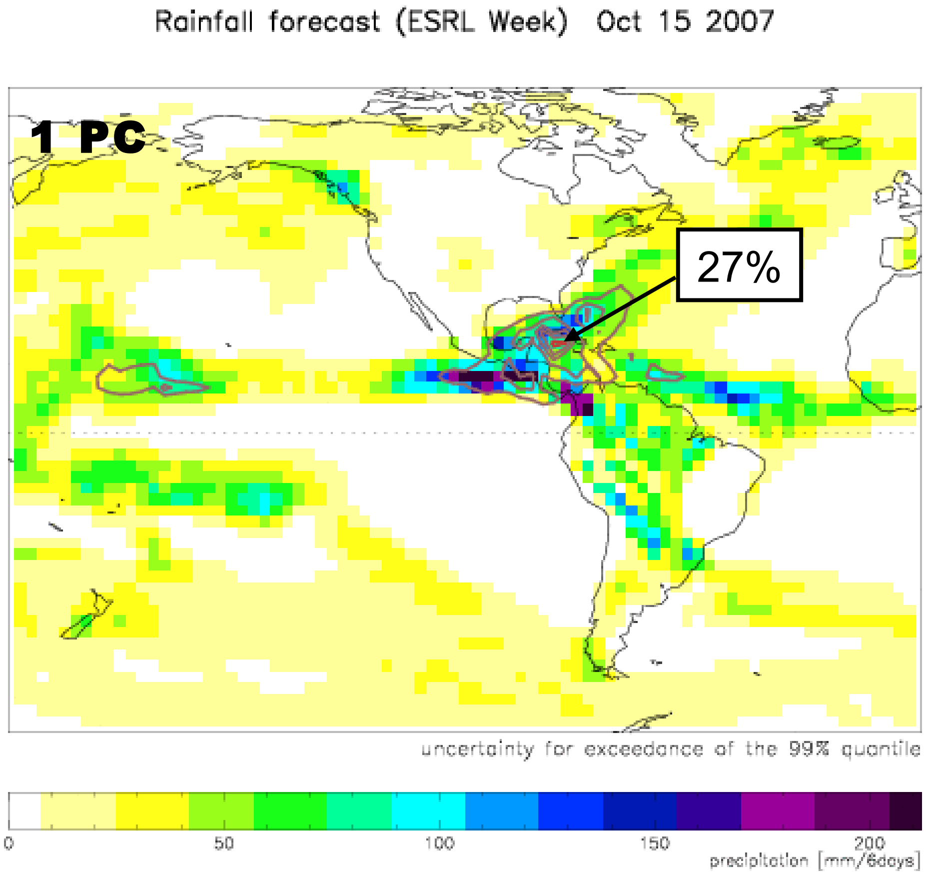
75°N

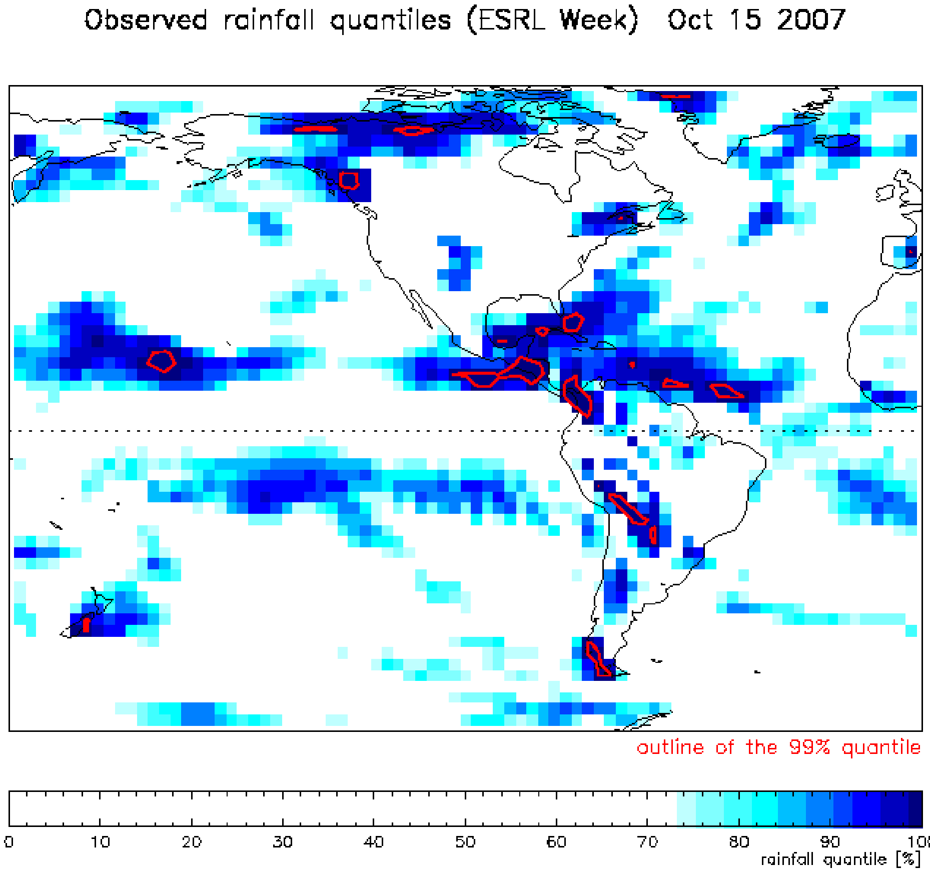


65°S

r_{max} = 2000 km · cos(latitude)

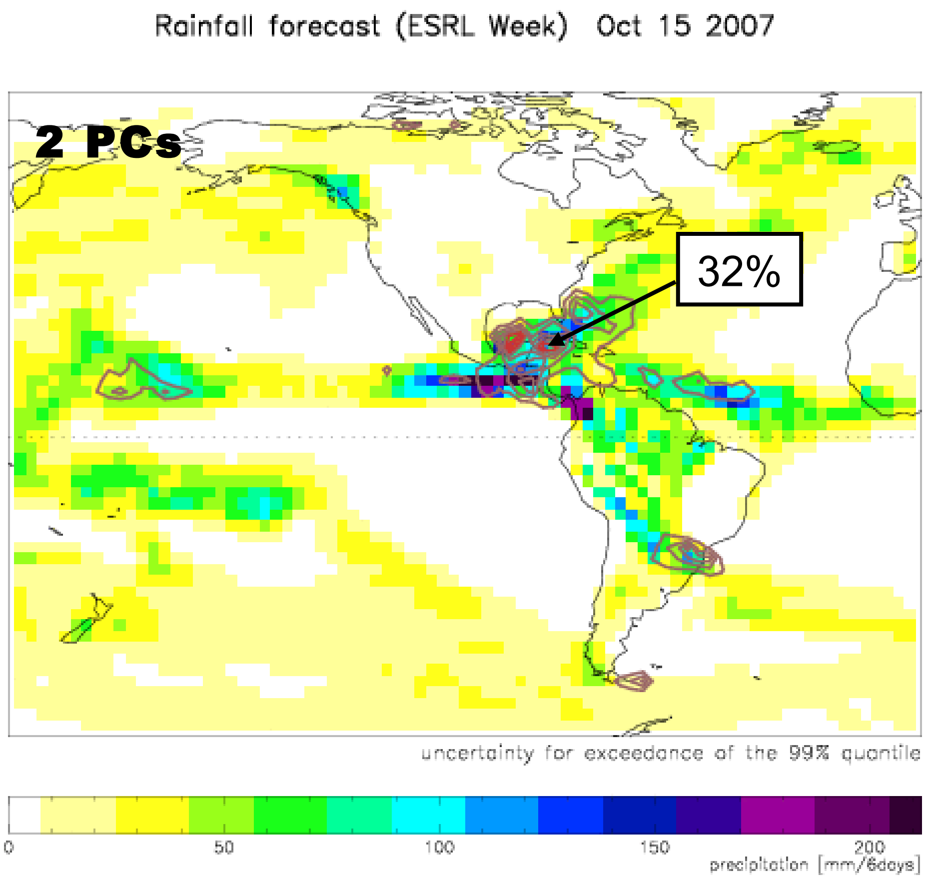
Probabilities for heavy precipitation (99% quantile)

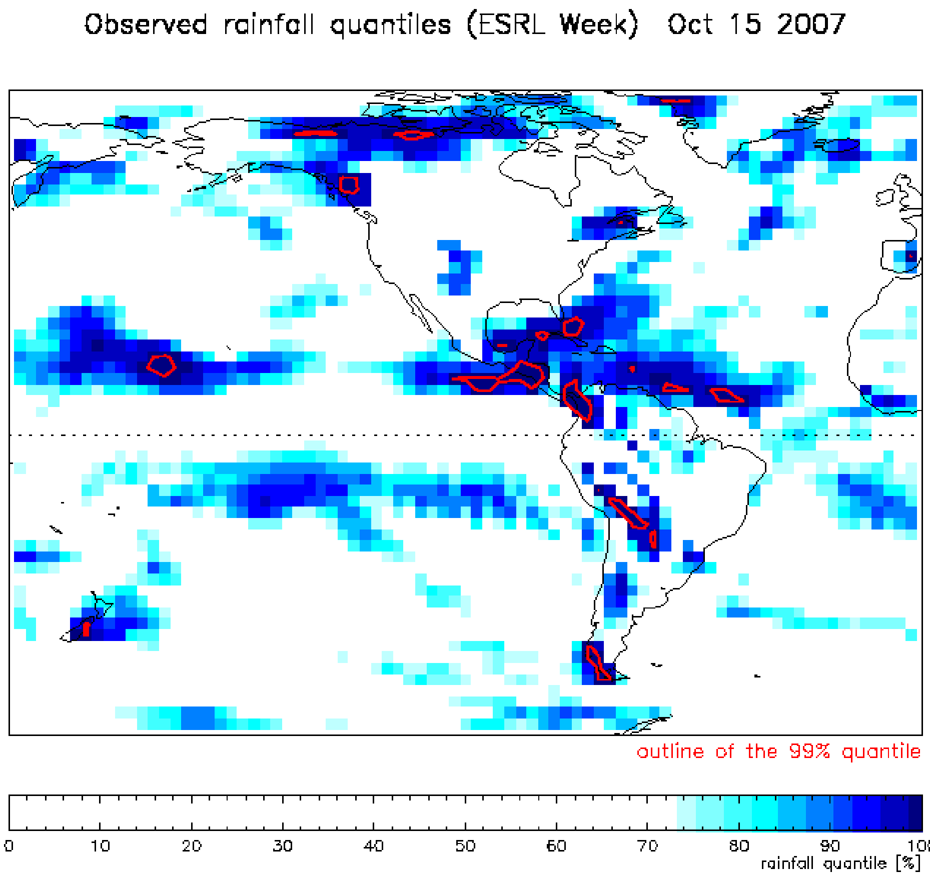




→Spatial rainfall patterns are reflected by principal components

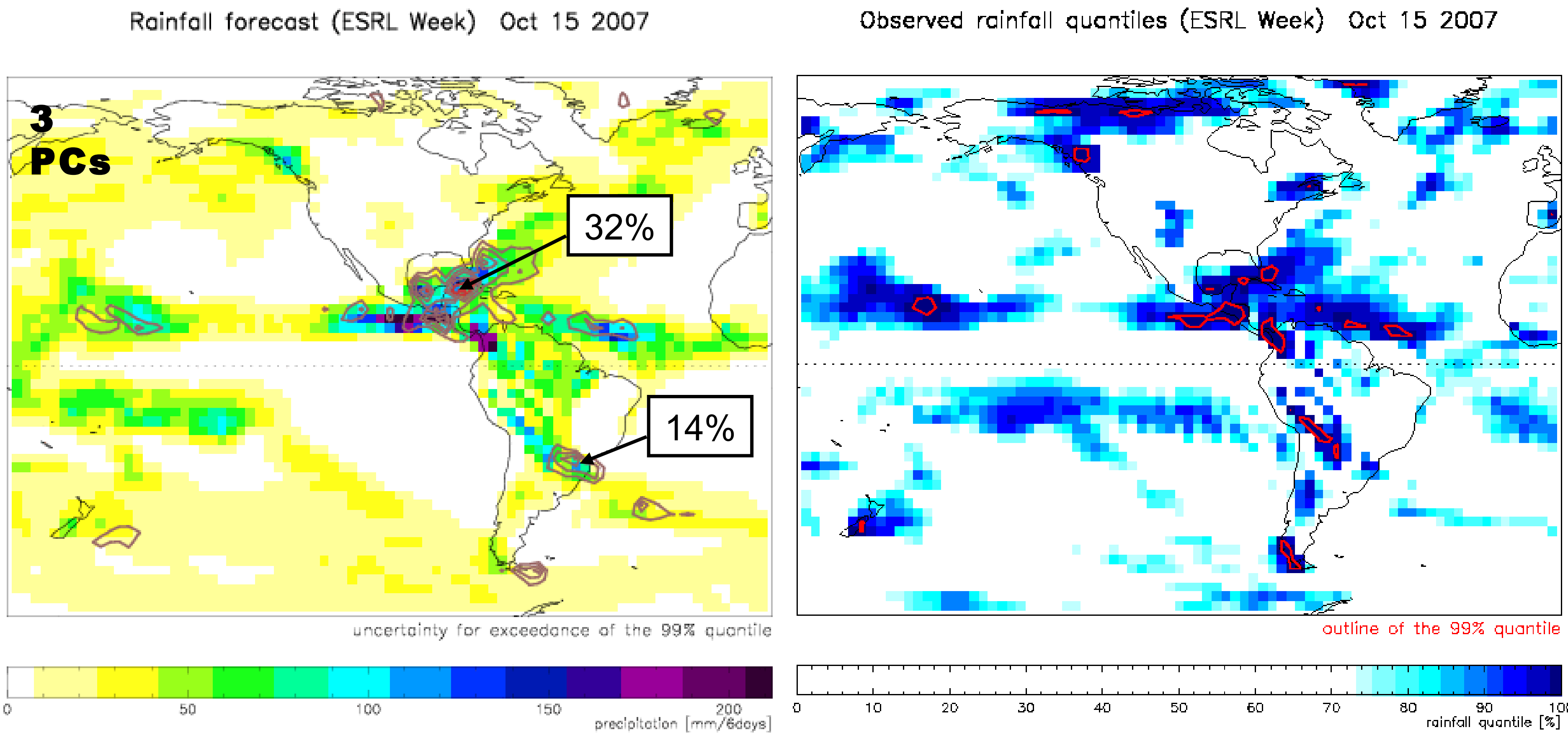
Probabilities for heavy precipitation (99% quantile)





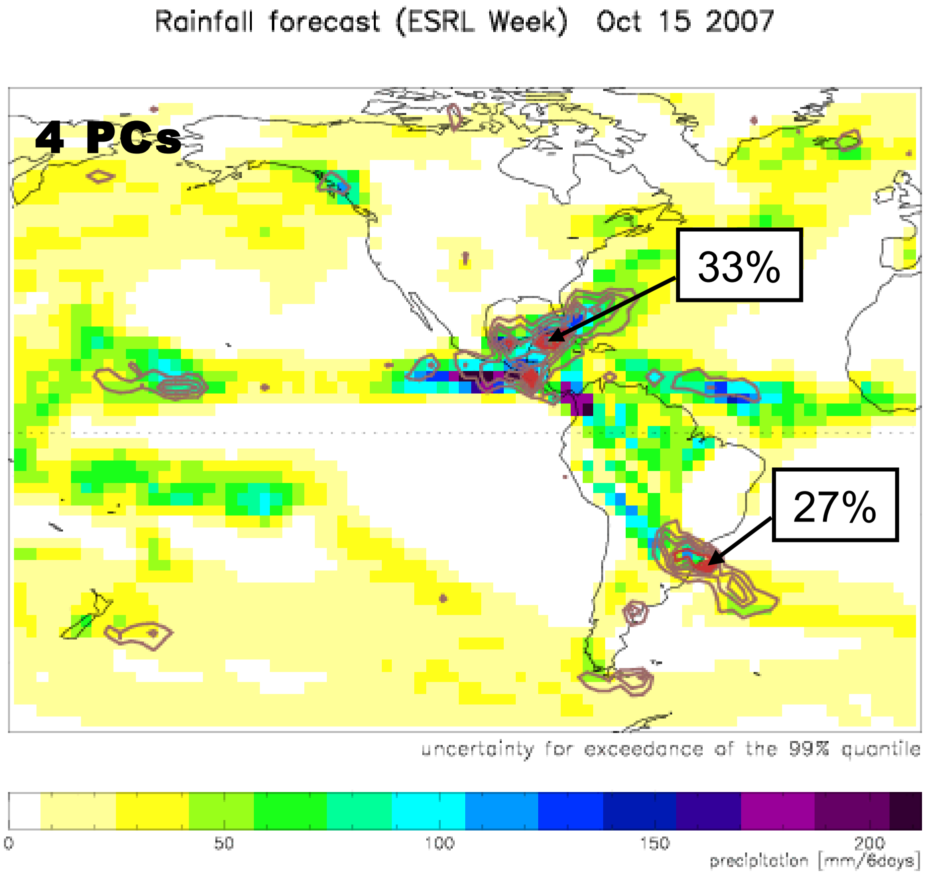
→Spatial rainfall patterns are reflected by principal components

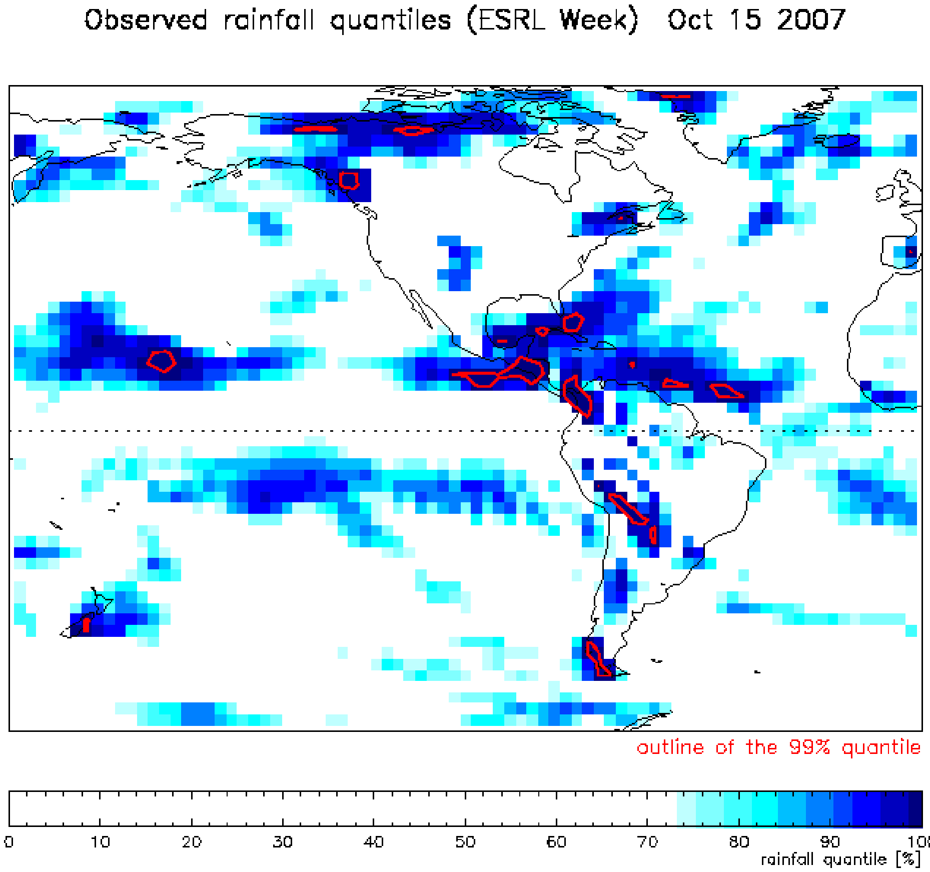
Probabilities for heavy precipitation (99% quantile)



→Spatial rainfall patterns are reflected by principal components

Probabilities for heavy precipitation (99% quantile)

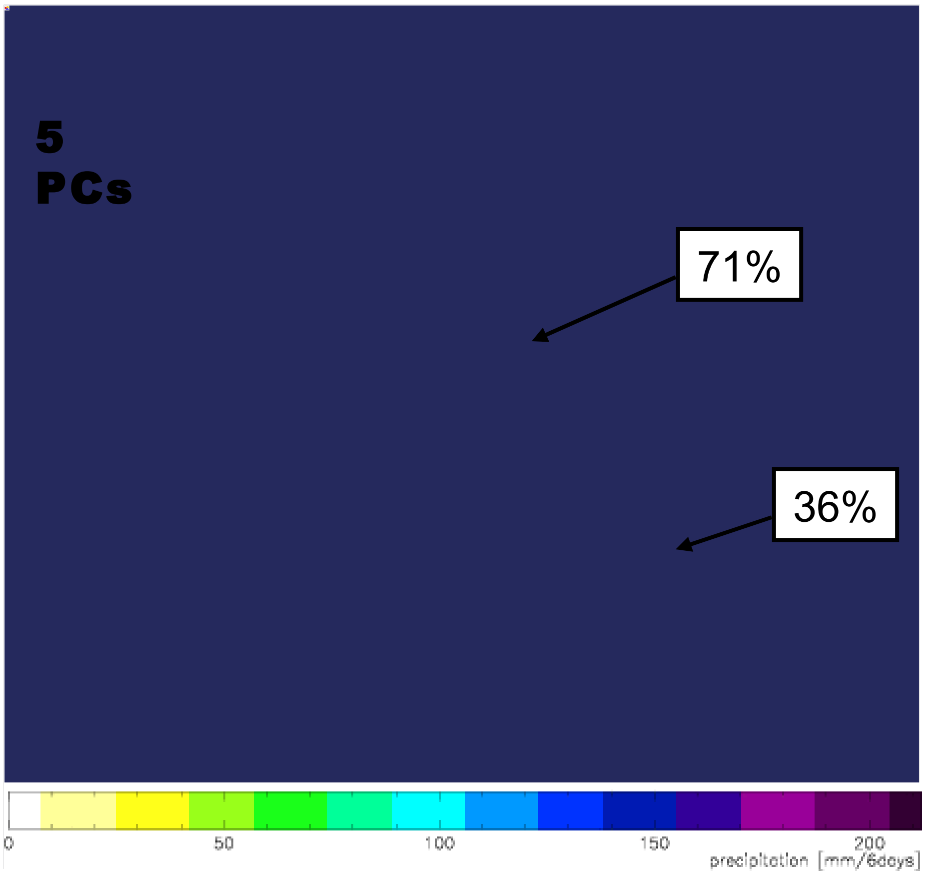


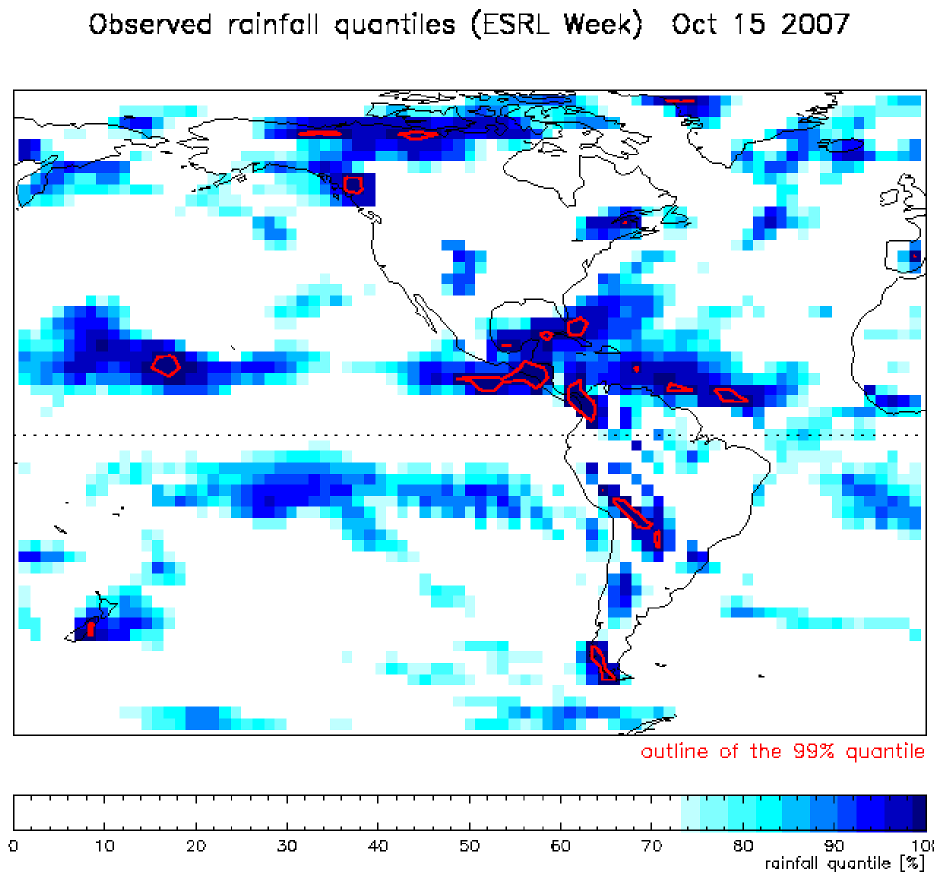


→Spatial rainfall patterns are reflected by principal components

Postprocessing methodology

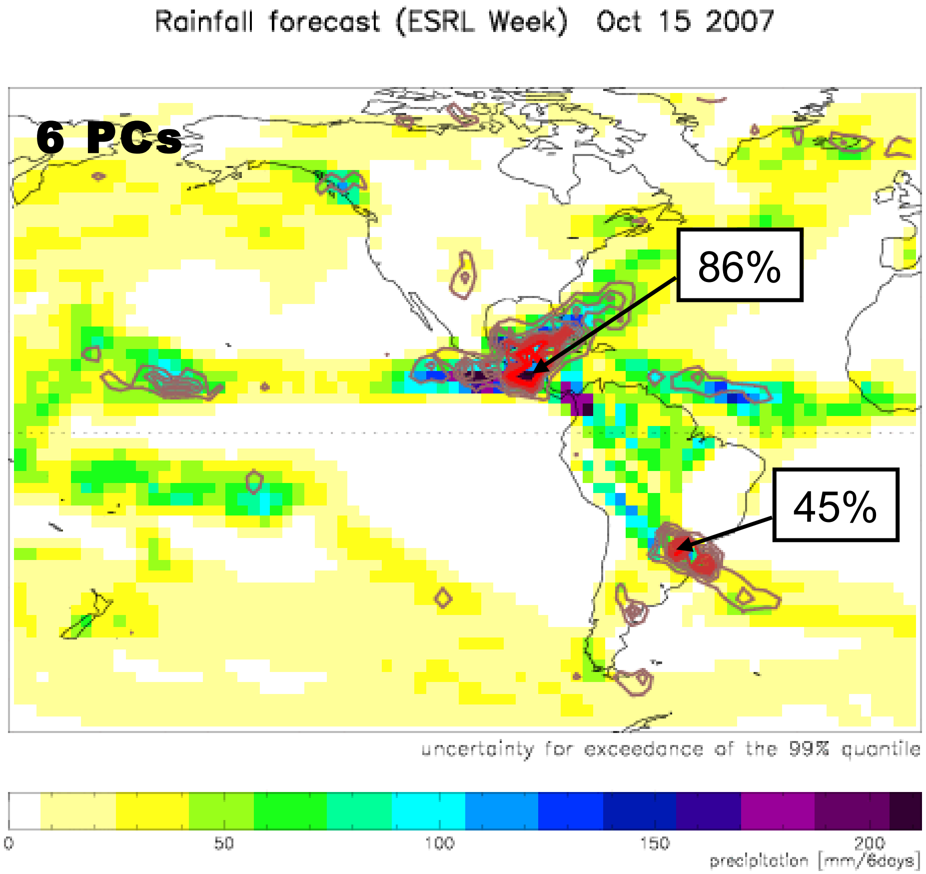
Probabilities for heavy precipitation (99% quantile)

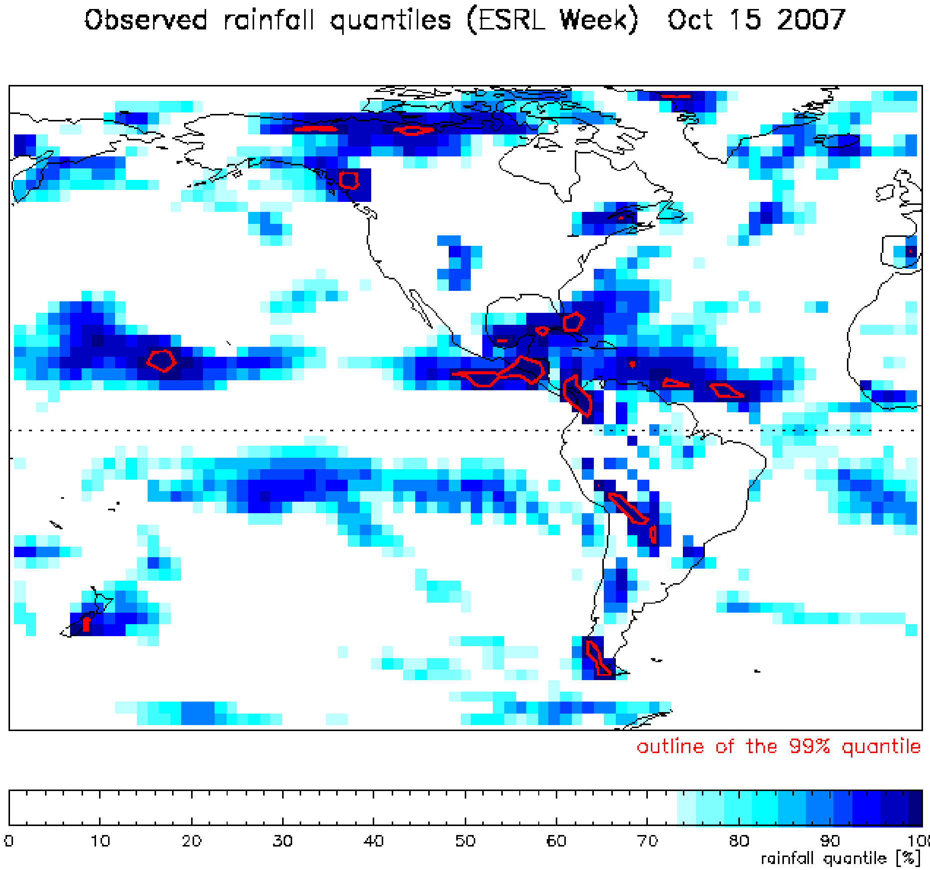




→Spatial rainfall patterns are reflected by principal components

Probabilities for heavy precipitation (99% quantile)



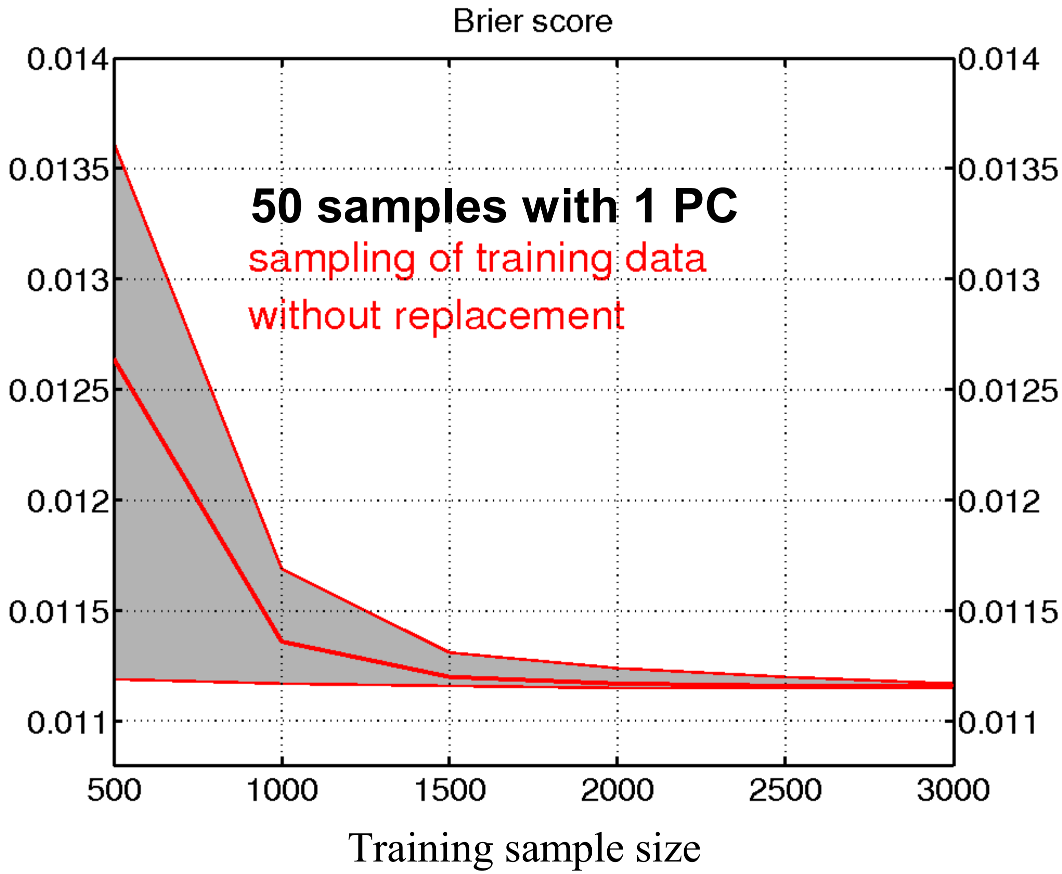


→Spatial rainfall patterns are reflected by principal components

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Sample sensitivities



99% observed quantile

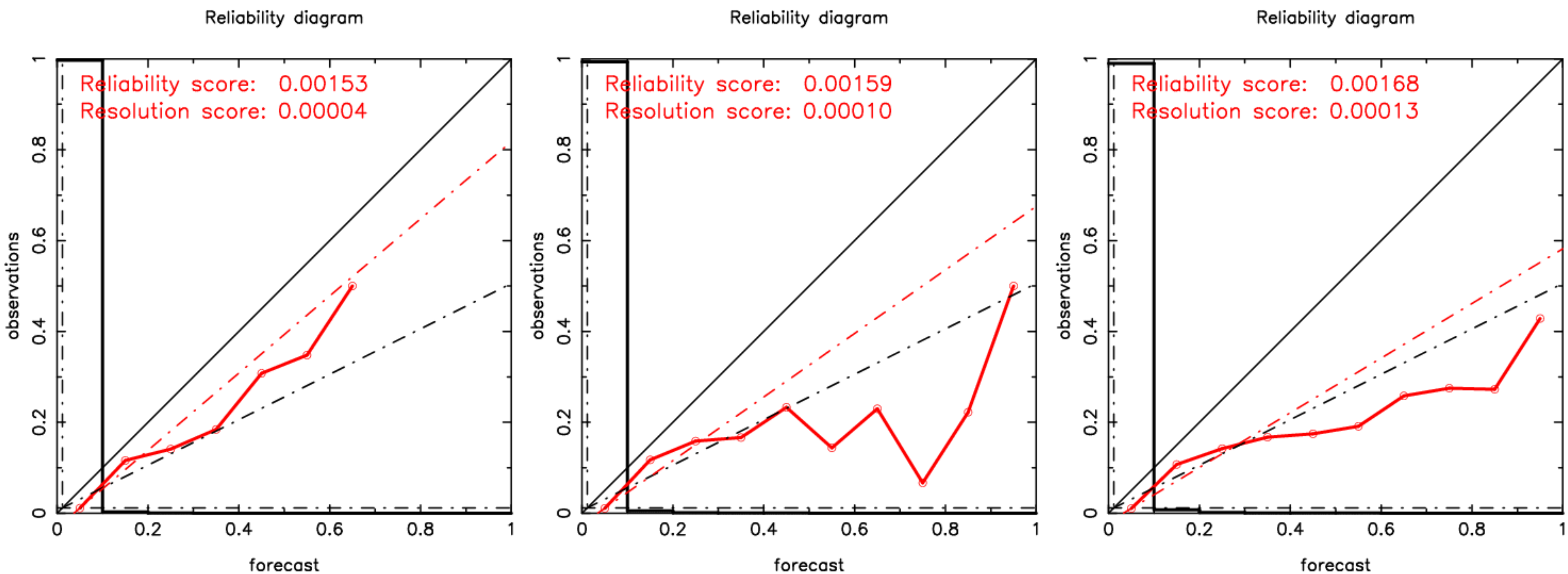
- **Saturation** for large training samples
- No significant im-provement for sample sizes over 1500 days

→ Improvements are due to a lower **reliability component**.
→ **Multiple principal components** display a much poorer performance for small sample sizes.

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Statistical sensitivities



1PC 3PCs 5PCs

→Risk of **overfitting** by use of too many PCs

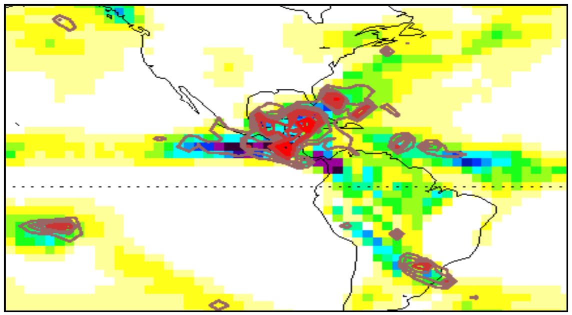
→Slight gain of **sharpness** at the expense of **reliability**

Overview

- Introduction of datasets
- Verification of direct model output
- Postprocessing methodology
- Sample sensitivities
  - Sample size of training data
- Statistical sensitivities
  - Number of principal components
- Summary and outlook

Summary and outlook

1.) Introduction of calibration methodology



2.) Probabilistic verification

→Best performance for just 1 PC

→Resolution counteracts with reliability

3.) Better tuning of methodology needed

→Test of different neighborhood sizes

→Selective use of multiple PCs

→Identification of highest skillful quantile

4.) Quantile estimation with logistic regression